Asian Herpetological Research 2016, 7(1): 1–11
DOI: 10.16373/j.cnki.ahr.150053

ORIGINAL ARTICLE

Allelic Polymorphism, Gene Duplication and Balancing Selection of MHC Class IIB Genes in the Omei Treefrog (*Rhacophorus omeimontis*)

Li HUANG[#], Mian ZHAO[#], Zhenhua LUO and Hua WU[*]

Institute of Evolution and Ecology, *International Research Centre of Ecology and Environment*, *School of Life Sciences*, *Central China Normal University*, 152 *Luoyulu*, *Hongshan District*, *Wuhan* 430079, *China*

Abstract The worldwide declines in amphibian populations have largely been caused by infectious fungi and bacteria. Given that vertebrate immunity against these extracellular pathogens is primarily functioned by the major histocompatibility complex (MHC) class II molecules, the characterization and the evolution of amphibian MHC class II genes have attracted increasing attention. The polymorphism of MHC class II genes was found to be correlated with susceptibility to fungal pathogens in many amphibian species, suggesting the importance of studies on MHC class II genes for amphibians. However, such studies on MHC class II gene evolution have rarely been conducted on amphibians in China. In this study, we chose Omei treefrog (*Rhacophorus omeimontis*), which lived moist environments easy for breeding bacteria, to study the polymorphism of its MHC class II genes and the underlying evolutionary mechanisms. We amplified the entire MHC class IIB exon 2 sequence in the *R. omeimontis* using newly designed primers. We detected 102 putative alleles in 146 individuals. The number of alleles per individual ranged from one to seven, indicating that there are at least four loci containing MHC class IIB genes in *R. omeimontis*. The allelic polymorphism estimated from the 102 alleles in *R. omeimontis* was not high compared to that estimated in other anuran species. No significant gene recombination was detected in the 102 MHC class IIB exon 2 sequences. In contrast, both gene duplication and balancing selection greatly contributed to the variability in MHC class IIB exon 2 sequences of *R. omeimontis*. This study lays the groundwork for the future researches to comprehensively analyze the evolution of amphibian MHC genes and to assess the role of MHC gene polymorphisms in resistance against extracellular pathogens for amphibians in China.

Keywords MHC class IIB, polymorphism, gene duplication, balancing selection, *Rhacophorus omeimontis*

1. Introduction

Proteins encoded by genes of the major histocompatibility complex (MHC) play an important role in the adaptive immunity of jawed vertebrates (Bernatchez and Landry, 2003). Two major groups of these proteins are highly concerned: MHC class I and class II. MHC class I proteins are expressed on the surface of nearly all nucleated somatic cells where they recognize intracellular pathogens such as viruses and present them to $CD8^+$ T-cells, while MHC class II molecules are primarily expressed on certain immune cells and present extracellular pathogens such as bacteria to $CD4^+$ lymphocytes (Hughes and Yeager, 1998; Bernatchez and Landry, 2003; Sommer, 2005). The MHC class II molecule is a heterodimer that consists of an α-chain and a β-chain. Each chain contains a cytoplasmic tail, a transmembrane domain, and extracellular domains 1 and 2. The domain 1, which is primarily encoded by exon 2 of MHC class II genes, is the key region responsible for the recognition of antigens (Hughes and Nei, 1990; Hughes and Yeager, 1998, Piertney and Oliver, 2006).

[#] These authors contributed equally to this work.
[*] Corresponding author: Prof. Hua WU, from Central China Normal University, Wuhan, China, with his research focusing on molecular ecology of amphibians.
E-mail: wuhua@mail.ccnu.edu.cn
Received: 28 July 2015 Accepted: 20 November 2015

Specifically, domain 1 from chains α and β (α1 and β1) form peptide-binding groove where antigen binding sites (ABSs) are situated according to the crystal structure of human MHC class II molecules (Brown *et al.*, 1993; Tong *et al.*, 2006).

High level of polymorphisms is commonly found in MHC genes, especially in ABSs (Bernatchez and Landry, 2003; Piertney and Oliver, 2006). Three major mechanisms have been proposed to explain the high polymorphisms of MHC genes: gene duplication (Baker *et al.*, 2006; Bryja *et al*., 2006; Balakrishnan *et al*., 2010; Bollmer *et al*., 2010; Sato *et al.*, 2011; Xu *et al*., 2011; Kiemnec-Tyburczy *et al.*, 2012), gene recombination (Reusch and Langefors, 2005; Schaschl *et al.*, 2006; Xu *et al.*, 2010; Smith *et al.*, 2011), and balancing selection (Aguilar *et al.*, 2004; Kamath and Getz, 2011). Gene duplication increases the prevalence of duplicated genes, providing additional genetic material on which abundant variations accumulate. Some duplicate genes are maintained in the genome for a long time, executing their original function, whereas others might be deleted or become dysfunctional due to deleterious mutations (Nei *et al.*, 1997; Nei and Rooney, 2005). Recombination within or between loci reshuffles the MHC sequences, and this process also increase sequence variability (She *et al*., 1991; Begovich *et al*., 1992; Andersson and Mikko, 1995; Jakobsen *et al.*, 1998; Zhao *et al*., 2013). Balancing selection refers to selection processes under which massive alleles are maintained in the gene pool of a population for a very long time (Bernatchez and Landry, 2003). Regardless of the exact mechanisms underlying MHC variation, the level of MHC polymorphism and the prevalence of certain MHC alleles were both found to be correlated with the ability of an organism or a population to fight disease (Edwards and Hedrick, 1998; Jeffery and Bangham, 2000; Bernatchez and Landry, 2003; Spurgin and Richardson, 2010). Therefore, species or populations displaying low MHC variability have been suggested to be particularly vulnerable to disease and, consequently, at a higher risk of extinction (O'Brien and Evermann, 1988; Sommer, 2005; Siddle *et al*., 2007).

Amphibians have undergone catastrophic declines worldwide (Stuart *et al*., 2004; Lips *et al*., 2006; Pounds *et al*., 2006). Infectious diseases, particularly chytridiomycosis caused by *Batrachochytrium dndrobatidis* (Bd), have been considered as one of the most important factors contributing to the declines in amphibians in many regions (Berger *et al*., 1998; Garner *et al*., 2006; Lips *et al*., 2006; Pounds *et al*., 2006; Morgan *et al*., 2007). Because MHC II molecules are responsible for immunity against fungi, increasing researches have focused on the polymorphisms and the evolution of MHC class II genes in amphibians (Bos and DeWoody, 2005; Hauswaldt *et al*., 2007; Babik *et al*., 2008, 2009; May and Beebee, 2009; Zeisset and Beebee, 2009, 2013; Kiemnec-Tyburczy *et al.*, 2010, 2012; Lillie *et al.*, 2015). The number of loci varies among amphibian species, and genetic diversity is commonly high (Laurens *et al.*, 2001; Bos and DeWoody, 2005; Hauswaldt *et al*., 2007; Babik *et al*., 2008, 2009; Zeisset and Beebee, 2013; Lillie *et al.*, 2015). Gene recombination (Babik *et al*., 2008, 2009; Lillie *et al.*, 2015) and balancing selection (Bos and DeWoody, 2005; Hauswaldt *et al*., 2007; Babik *et al*., 2008, 2009; Lillie *et al.*, 2015) have been found to contribute to polymorphisms within MHC class II genes in many amphibian species. Moreover, correlations between MHC class II genes and resistance or susceptibility to fungi have been found in amphibians (May *et al.*, 2011; Savage and Zamudio, 2011; Bataille *et al.*, 2015).

In contrast to the relatively vast number of studies on fungal infection and MHC class II gene evolution in amphibians native to other countries (Babik *et al*., 2008, 2009; Kiemnec-Tyburczy *et al.*, 2010; Baláž *et al.*, 2014; Bataille *et al.*, 2015), very few similar investigations have been conducted on amphibians in China (Bai *et al.*, 2010; Wei *et al.*, 2010; Li *et al.*, 2012; Shu *et al.*, 2013; Zhu *et al.*, 2013; Yu *et al.*, 2014). Although such research in China is limited, one study revealed that several native amphibians in China were infected by Bd (Bai *et al*., 2010), and another study found that two MHC DAB alleles were significantly associated with resistance to *Aeromonas hydrophila* (Yu *et al.*, 2014). Therefore, further studies are needed to assess the effects of fungi and bacteria on the demographics of amphibians native to China and to explore the role of MHC class II genes in these effects.

The Omei treefrog (*Rhacophorus omeimontis*), which belongs to the Rhacophoridae family of amphibians, is primarily distributed in wooded, moist mountainous regions of southwestern China at altitudes ranging from 700 to 2000 m (Zhao and Adler, 1993; Fei and Ye, 2001). *R. omeimontis* breeds on tree branches over static pools, where their larvae grow and develop. These habitats expose *R. omeimontis* to large amounts of microorganisms, which imposes considerable selection pressure on immune-related MHC genes. Our previous study reported a high genetic diversity in MHC class I genes of *R. omeimontis* (Zhao *et al.*, 2013) and suggested that pathogen-mediated balancing selection played an

important role in the formation and maintenance of the high MHC polymorphisms. However, MHC class II genes that are associated with immunity against fungi and bacteria have not been characterized in this species. Here, we 1) designed primers to isolate the entire exon 2 region of MHC IIB genes in *R. omeimontis*, 2) characterized the MHC IIB genes in *R. omeimontis*, and 3) performed a preliminary investigation of the mechanisms underlying the formation and maintenance of MHC gene polymorphisms.

2. Materials and Methods

2.1 Sampling and nucleic acid extraction During the breeding period of *R. omeimontis* in 2011 (from mid-April to mid-June), we collected 146 toe samples from the *R. omeimontis* population in Fengtongzhai, Baoxing County, Sichuan Province, using the toe clipping method. Specifically, we searched for *R. omeimontis* near the static pools and captured them by hand. Noticeably, if the frogs were copulating, we would not capture them until the frogs finished the breeding and left the spawning group, in order to minimize the disturbance to their breeding behaviors. A small piece of toe was removed from the hind foot of each captured frogs using surgical scissors. Antiseptic was used on both the scissors and the frog toes. The frogs were then released back to the site of capture as soon as possible. All toe samples were preserved in 99.5% ethanol. In addition, we captured three live *R. omeimontis* individuals (one from Fengtongzhai, Sichuan Province; two from Badagongshan, Hunan Province) for the following RNA extraction. Genomic DNA was extracted from the toe clips using the TIANamp Genomic DNA Kit (TIANGEN, Beijing, China). The three live frogs were sacrificed, and their liver tissue was excised. Total RNA was then isolated from the fresh liver tissue using TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. The quality and the concentration of the total DNA/RNA samples were assessed using agarose gel electrophoresis and ultraviolet spectrophotometric analysis. First-strand complementary DNA (cDNA) was synthesized from the total RNA samples using the reverse transcriptase M-MLV (TaKaRa, Dalian, China) and oligo$(dT)_{18}$ primers. Both the genomic DNA (gDNA) and the cDNA were stored at –20°C for future use.

2.2 Ethics statements Frog sampling was conducted under ethical approval granted by Central China Normal University and the Management Bureau of the Badagongshan/Fengtongzhai National Nature Reserve. No significant adverse stress was imposed to *R. omeimontis*.

2.3 Primer design We designed primers to isolate the entire exon 2 of MHC class IIB genes from the gDNA of *R. omeimontis*. All of the new primers in this study were designed using Primer3 (Rozen and Skaletsky, 2000). First, we searched NCBI for published MHC class IIB alleles. The sequences from *Xenopus laevis* (D13684, D13688, D13687, and D13685), *Rana pipiens* (HQ025936), *R. sylvatica* (HQ025938), *R. clamitans* (HQ025932), *R. palustris* (HQ025934), *R. catecbeiana* (BT081564), *Epidalea calamita* (HQ388291), *Ambystoma mexicanum* (AF209117 and AF209115), *A. tigrinum* (DQ125479 and DQ125480), *Mus musculus* (NM_207105), and *Homo sapiens* (NM_002124) were compiled and aligned. Based on the conserved regions of these sequences, we designed one set of degenerate primers, MHCIIB-4F and MHCIIB-4R (Table 1). We amplified sequences from cDNA using these two new primers and obtained the entire exon 3 (282 bp) and partial regions of exon 2 (35 bp) and exon 4 (11 bp). Second, we used the conserved regions of these newly amplified sequences to design another primer (II 5-NGSP; Table 1). Using the new primer II 5-NGSP and the existing primer MHC-F (Hauswaldt *et al.*, 2007), we isolated a large region of MHC IIB exon 2 (254 bp) from the cDNA of *R. omeimontis*. Third, using the sequence information on the known portion of exon 2, we designed four primers (RB-0a, RB-1a, RB-1ac, and RB-2a; Figure 1 and Table 1) that corresponded to four previously published primers (LB-0a, LB-1a, LB-1ac, and LB-2a, respectively) (Liu and Chen, 2007). Then, we conducted thermal asymmetric interlaced PCR (TAIL

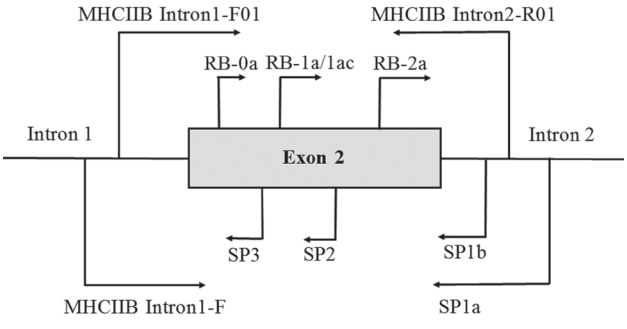


Figure 1 Illustration of the locations of the PCR oligonucleotide primers that were used to amplify exon 2 of MHC class IIB genes in *R. omeimontis*. Two sets of primers (RB-0a, RB-1a, RB-1ac, and RB-2a; SP1a, SP1b, SP2, and SP3) were used for TAIL PCR and gene walking, respectively. Finally, two pairs of newly designed primers (MHCIIB Intron1-F and SP1a; MHCIIB Intron1-F01 and MHCIIB Intron2-R01) were adopted to amplify the entire MHC IIB exon 2 in genomic DNA from 146 treefrog individuals.

PCR) using the four sets of primers with gDNA as the template to isolate the downstream intron 2 sequences. Fourth, we designed four primers (SP1a, SP1b, SP2, SP3; Figure 1 and Table 1) based on the newly obtained partial exon 2 and intron 2 sequences. Combining with the existing primers in gene walking kits (AP1, AP2, AP3, and AP4), we applied the gene walking PCR method to the gDNA to amplify the upstream intron 1 sequences. Finally, based on the flanking intronic sequences of exon 2, two pairs of primers (MHCIIB Intron1-F and SP1a; MHCIIB Intron1-F01 and MHCIIB Intron2-R01; Figure 1 and Table 1) were developed to amplify the entire exon 2 of MHC IIB genes from genomic DNA of *R. omeimontis*. The optimum annealing temperatures of the two set of primers were 60 °C and 62 °C, respectively.

2.4 DNA amplification, cloning and sequencing The newly designed primers were used to amplify the entire MHC IIB exon 2 sequences in gDNA from the 146 toe samples. PCR was conducted under the following conditions: 95 °C for 5 min; 30 cycles of 98 °C for 10 s, 60 °C or 62 °C for 30 s, and 72 °C for 45 s; and a final extension at 72 °C for 10 min. The PCR product was purified using a Universal DNA Purification Kit (TIANGEN) and was then cloned into the pMD19-T vector (TaKaRa). The recombinant DNA was then transformed into *Escherichia coli* DH5α competent cells. Ten positive clones per individual were selected via PCR using the M13F and M13R primers and were prepared for sequencing.

2.5 Sequence analyses The sequences we obtained were edited using Chromas version 2.22 (Technelysium Ltd, Helensvale, Queensland, Australia) and aligned using the ClustalW module in MEGA v6.0 (Tamura *et al*., 2013). Exon 2 sequences of MHC IIB genes from other *R. omeimontis* populations (unpublished data in our laboratory) were included in the alignment. We defined alleles as the sequences found in more than one *R. omeimontis* individual.

Polymorphism analyses of the alleles were conducted using DnaSP v5.0 software (Librado and Rozas, 2009). The overall mean distances of nucleotides (Kimura-2-parameter model, K2P) and amino acids (the Poisson model) among the alleles were both computed using MEGA v6.0 (Tamura *et al*., 2013). Standard errors of the estimates were projected using 1000 bootstrap replicates. The ABSs on the alleles were identified according to the codons involved in pathogen binding on the β chain of MHC class II molecules in humans (Brown *et al*., 1993; Tong *et al*., 2006). The number of synonymous substitutions per synonymous site (dS) and the number of non-synonymous substitutions per non-synonymous site (dN) in different regions of the alleles were calculated using MEGA v6.0 (Tamura *et al*., 2013) according to the Nei-Gojobori method with the Jukes-Cantor correction (Nei and Gojobori, 1986). Natural selection can be tested using the ratio of the nonsynonymous to synonymous substitution rates (dN/dS = ω) (Hughes and Nei, 1989; Garrigan and Hedrick, 2003). A codon-based Z test for selection, which is implemented in MEGA, was conducted to test for historical signals of positive selection in MHC class II alleles.

Gene recombination was screened using the Recombination Detection Program (RDP) version 4.50 (Martin *et al*., 2015). Recombination was considered to have occurred when positive signals were found in more than half of the applied methods (RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiSscan, PhylPro, LARD, and 3Seq). Positively selected sites were first detected at Datamonkey website (http://www.datamonkey.org) using two models, fixed effects likelihood (FEL) (Pond and Frost, 2005) and mixed-effects model of evolution (MEME) (Murrell *et al*., 2012). Subsequently, the codeml subroutine contained in PAML v4.0 (Yang, 1997; Yang, 2007) was also applied to test for positive selection in MHC class IIB genes. Variable ω (dN/dS) among codons was estimated using the maximum likelihood method (Goldman and Yang, 1994; Yang *et al*., 2000). The following four models of selection were tested: M1a (for nearly neutral evolution), M2a (for positive selection), M7 (for nearly neutral evolution in which the beta distribution approximated the ω variation), and M8 (for positive selection in which the beta distribution approximated the ω variation) (Goldman and Yang, 1994; Yang *et al*., 2000; Yang *et al*., 2005). The likelihood ratio test (LRT) was used to compare the neutral models (M1a and M7) with the nested models that allowed for positive selection (M2a and M8). Sites under positive selection were identified in models M2a and M8 using the Bayes empirical Bayes (BEB) procedure (Yang *et al*., 2005).

3. Results

3.1 Isolation of MHC IIB exon 2 In this study, we designed primers that effectively amplified the entire MHC IIB exon 2 sequences in *R. omeimontis*. Previous studies of MHC class II genes primarily focused on other species, such as *X. laevis* and those of the *Rana* genus (Sato *et al*., 1993; Ohta *et al*., 2006; Zeisset and Beebee, 2009; Kiemnec-Tyburczy *et al*., 2010). Moreover, most

of these studies isolated only a partial exon 2 region of MHC class IIB genes (Hauswaldt *et al.*, 2007; Zeisset and Beebee, 2009). Consequently, when we compiled the published MHC class II genes to identify conserved reference sequences for primer design, we found that the available data were primarily comprised of partial exon 2 sequences in species that are not phylogenetically close to *R. omeimontis*. Using primers designed from these sequences, we could initially amplify only a small region of MHC class IIB exon 2 sequences from cDNA of the studied treefrogs. Subsequently, the upstream and downstream sequences covering part of the intronic sequences flanking exon 2 could be amplified from gDNA using TAIL PCR and gene walking. The primers used to isolate the entire exon 2 from gDNA were ultimately designed based on the introns flanking exon 2.

Using the newly designed primers, we successfully amplified the exon 2 sequences of MHC IIB genes from 146 *R. omeimontis* samples. A total of 1460 sequences were obtained, and each sequence contained 267 nucleotides. 102 MHC IIB exon 2 alleles (GenBank accession numbers: KT276411–KT276512) were identified. All of these 102 alleles could be translated into sequences of 89 amino acids, with no insertions, deletions or stop codons (Figure 2). A homology analysis using BLASTN revealed high similarity (up to 85%) between our MHC IIB alleles and the published anuran MHC IIB genes. Additionally, the key amino acid sites that form disulfide bridges in human HLA β chains were highly conserved in our newly identified alleles (Cys10 and Cys74; Figure 2). These results suggested that our newly designed primers successfully isolated MHC IIB exon 2 from gDNA of *R. omeimontis*. Moreover, the number of alleles per individual ranged from one to seven, suggesting that at least four loci containing MHC IIB exon 2 are present in the genome of *R. omeimontis*.

3.2 Characterization of MHC genes The newly identified MHC IIB exon 2 alleles varied at 104 nucleotide sites (39.0% of the 267 nucleotide sites) and 60 amino acid sites (67.4% of the 89 amino acid sites). The nucleotide diversity (π) was 0.038 ±0.003, and the average number of nucleotide differences (k) was 10.092. The mean Poisson-modeled amino acid distance and the mean K-2P nucleotide distance were 0.082 ± 0.017 and 0.039 ± 0.006, respectively.

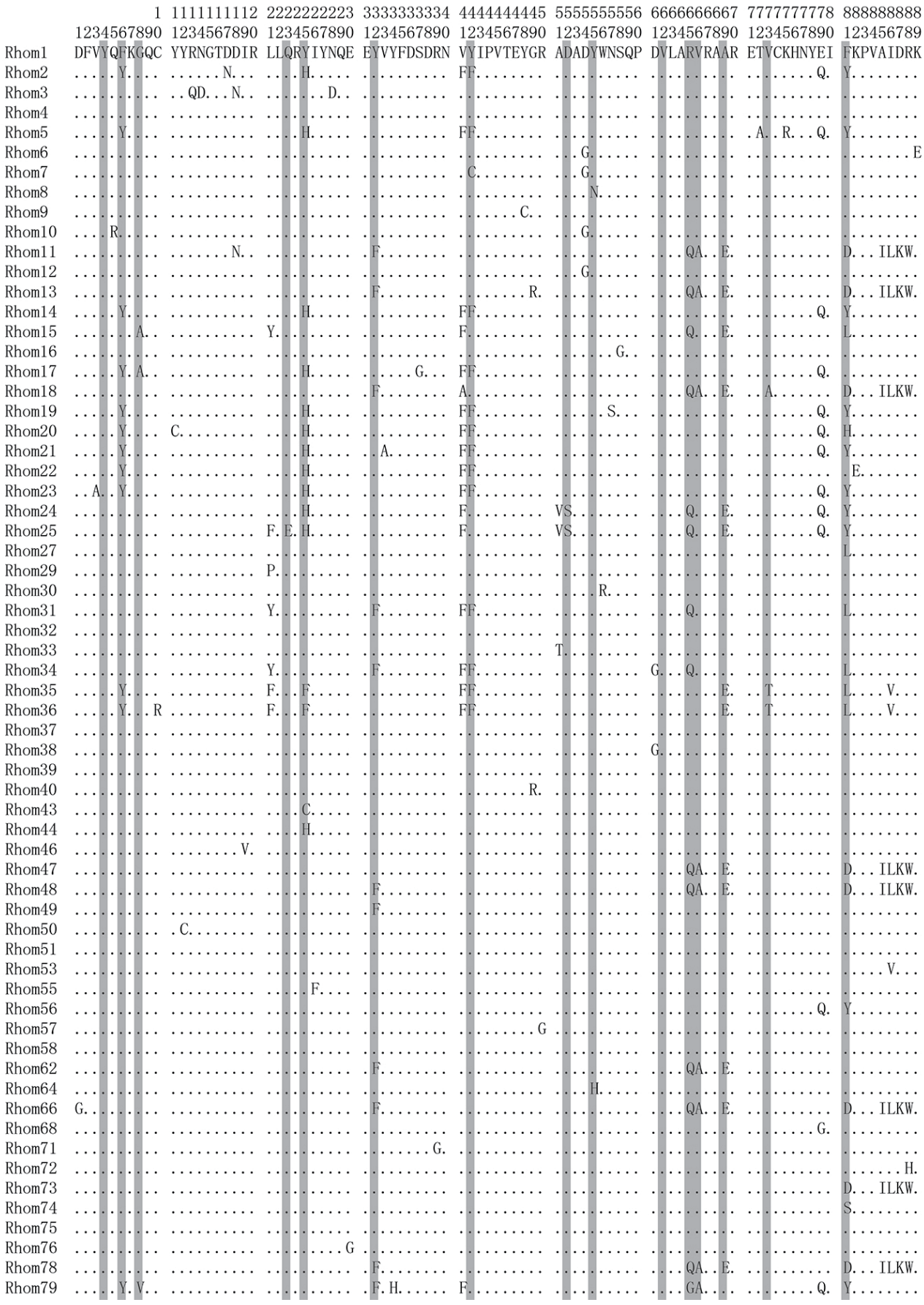
Fifteen putative ABSs were deduced according to the crystal structure model of the MHC class II β chain in humans (Brown *et al.* 1993; Tong *et al.* 2006). All estimated dN values were larger than the dS values for the entire region, the ABS region and the non-ABS region of our MHC IIB exon 2 sequences, with the highest dN/dS value observed for the ABS region. However, the differences between dN and dS were not significant for all of the three different regions (Table 2).

3.3 Detection of gene recombination and natural selection No signal of gene recombination was detected in RDP. Positive natural selection was then detected without considering gene recombination. Three positively selected sites (sites 6, 21 and 81) and one negatively selected site (site 60) were detected using FEL analysis (Table S1), and six sites (sites 6, 21, 52, 65, 73 and 81) were estimated to be under episodic diversifying selection based on MEME analysis (Table S2). Using the BEB method implemented in PAML, ten significantly positive selection sites were detected using models M2a and M8 (Table S3). Taken together, selection testing programs detected a total of twelve positively selected codons, nine of which matched peptide-binding sites predicted from the structures of human MHC class II molecules (Table 3; Figure 2).

4. Discussion

We designed intronic primers to amplify the entire MHC IIB exon 2 sequences in *R. omeimontis* for the first time. Evidence of our successful isolation of the MHC IIB alleles in *R. omeimontis* included the lack of indels and stop codons, the high homology of the obtained alleles with published anuran MHC IIB genes, and the conservation of amino acid sites relating to disulfide formation and antigen binding.

We isolated a total of 102 MHC IIB alleles from 146 *R. omeimontis* individuals. The allelic polymorphism of MHC IIB genes in *R. omeimontis* was not high compared to that in other amphibian species. For instance, in the 74 alleles isolated from 121 *Leiopelma hochstetteri* individuals, the average nucleotide diversity and the average number of nucleotide differences were 0.091 and 19.6, respectively (Lillie *et al.*, 2015). For the 8 alleles obtained from 215 *Rana temporaria* individuals, the mean amino acid and nucleotide distances were 0.147 and 0.077, respectively (Zeisset and Beebee, 2009). The relatively low polymorphism in *R. omeimontis* MHC IIB genes might result from relaxed natural selection, may simply reflect low genetic diversity at the genome level, or both. Genetic diversity measured by neutral genetic marker was also low in this *R. omeimontis* population (our unpublished data), indicating the probable role of genetic drift in the relatively low polymorphism of *R. omeimontis* MHC genes.


1 1111111112 2222222223 3333333334 4444444445 5555555556 6666666667 7777777778 8888888888
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Rhom1 DFVYQFKGQC YYRNGTDDIR LLQRYIYNQE EYVYFDSDRN VYIPVTEYGR ADADYWNSQP DVLARVRAAR ETVCKHNYEI FKPVAIDRK
Rhom2 Rhom3 Rhom4 Rhom5 Rhom6 Rhom7 Rhom8 Rhom9 Rhom10 Rhom11 Rhom12 Rhom13 Rhom14 Rhom15 Rhom16 Rhom17 Rhom18 Rhom19 Rhom20 Rhom21 Rhom22 Rhom23 Rhom24 Rhom25 Rhom27 Rhom29 Rhom30 Rhom31 Rhom32 Rhom33 Rhom34 Rhom35 Rhom36 Rhom37 Rhom38 Rhom39 Rhom40 Rhom43 Rhom44 Rhom46 Rhom47 Rhom48 Rhom49 Rhom50 Rhom51 Rhom53 Rhom55 Rhom56 Rhom57 Rhom58 Rhom62 Rhom64 Rhom66 Rhom68 Rhom71 Rhom72 Rhom73 Rhom74 Rhom75 Rhom76 Rhom78 Rhom79

(Continued Figure 1)

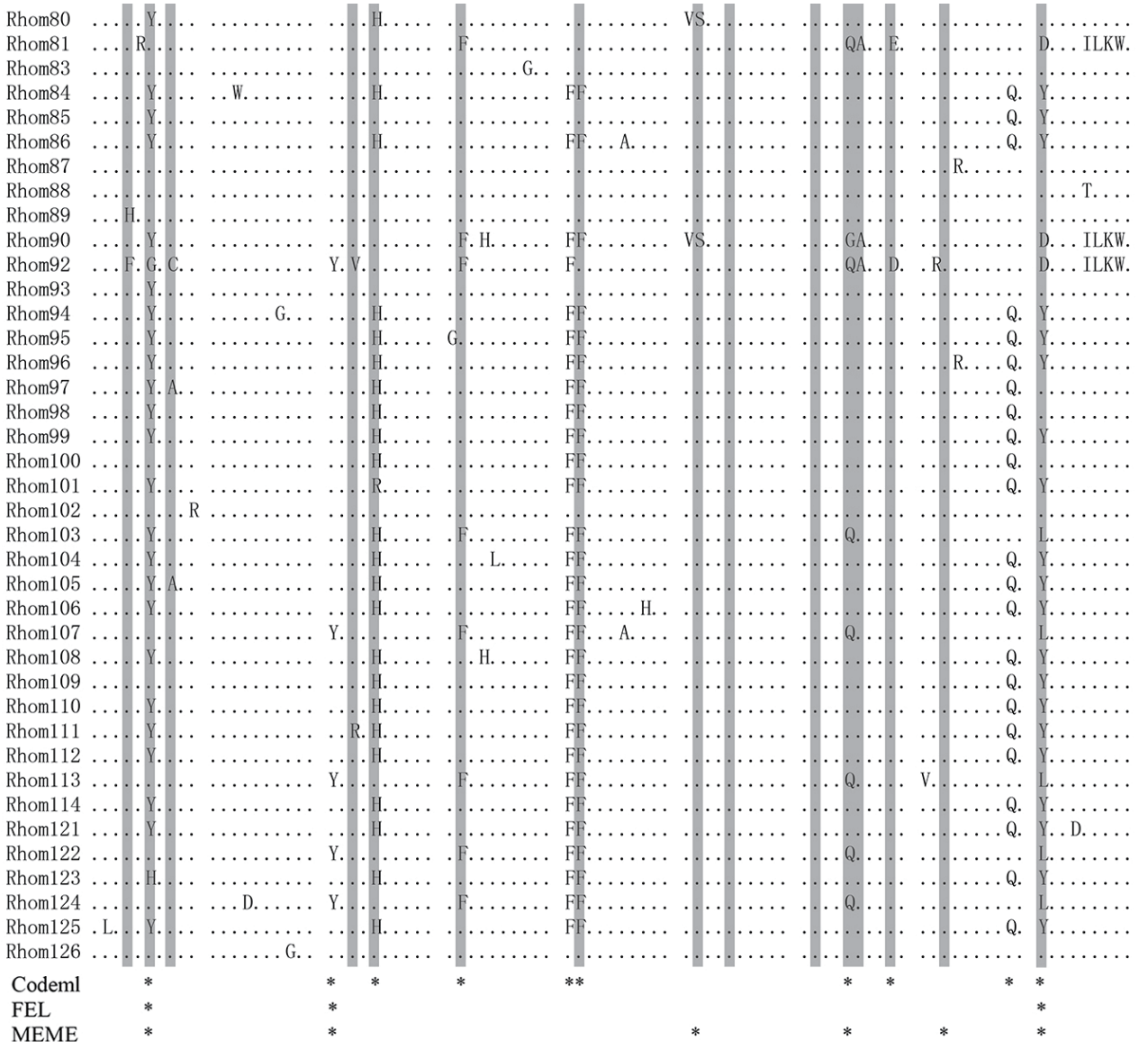


Figure 2 Sequence alignments of the 102 MHC IIB exon 2 alleles found in *R. omeimontis*. The dots in this figure show the sequence identity relative to the top sequence. The sites outlined in gray refer to the putative ABSs, which were deduced from human HLA molecules (Brown *et al.* 1993; Tong *et al.* 2006). The asterisks represent the codons subjected to positive selection based on inference of selection testing programs.

The number of alleles per individual ranged from one to seven, indicating that there were four loci of MHC class II genes in *R. omeimontis*. Given that the alleles were identified using conservative criteria and that rare alleles were easily missed in this study, the locus number was probably underestimated. Nevertheless, the richness of the MHC class II locus in *R. omeimontis* was relatively high among amphibian species. As shown in previous studies, a single MHC class II locus was found in *R. temporaria* (Zeisset and Beebee, 2009) and *A. tigrinum* (Laurens *et al*., 2001; Bos and DeWoody, 2005). Two loci have been detected in a broad range of amphibians, including *Xenopus* (Sato *et al.*, 1993; Ohta *et al.*, 2006), *Bufo* (May and Beebee, 2009; Zeisset and Beebee, 2013), *Rana* (Kiemnec-Tyburczy *et al.*, 2010), *Bombina bombina* (Hauswaldt *et al.*, 2007) and *Quasipaa spinosa* (Yu *et al*., 2014). In addition, four or more loci were found only in *Odorrana tormota* (Shu *et al*., 2013), *Mesotriton alpestris* (Babik *et al*., 2008) and *Triturus cristatus* (Babik *et al*., 2009). However, among these four or more loci, only one or two loci were commonly shown to be expressed (Babik *et al*., 2008, 2009). Given that the MHC IIB sequences

were isolated from gDNA in this study, the four loci in *R. omeimontis* might not be all expressed. Although the 102 exon 2 alleles appeared to be functional, some of these alleles were possibly pseudogenes because we had no data on the MHC gene regions outside of exon 2. This presumption is evidence-based considering the pattern of MHC IIB splice variants which were found in the Chinese giant salamander *Andrias davidianus* (Zhu *et al*., 2013). Compared with the full-length transcripts of the MHC IIB alleles, the shortened variants contained very similar β1 domains but lacked β2 domains (Zhu *et al*., 2013). Thus, it is difficult to determine whether the MHC alleles are functional if only part of the MHC gene sequences are available for analyses. However, regardless of whether the loci characterized in this study are effectively expressed, all of the alleles showed very high sequence similarity. These results suggested that the multiple loci containing MHC IIB gene sequences were generated via gene duplication (Nei *et al*., 1997; Nei and Rooney, 2005).

We did not detect significant gene recombination in the 102 alleles, and this result is consistent with previous studies on MHC class II genes in other anuran species (Bos and DeWoody, 2005; Kiemnec-Tyburczy *et al.*, 2010; Shu *et al.*, 2013; Bataille *et al.*, 2015). However, this finding may be due to our detection of a part region of MHC IIB alleles rather than the true absence of gene recombination in MHC IIB genes of *R. omeimontis*. To date, gene recombination has been detected in MHC IIB genes of many amphibians, including *L. hochstetteri* (Lillie *et al.*, 2015), *Q. spinosa* (Yu *et al*., 2014), *M. alpestris* (Babik *et al*., 2008) and *T. cristatus* (Babik *et al*., 2009). Further study should include the entire region and increase the number of MHC class II genes analyzed to more accurately understand the contribution of gene recombination to the evolution of MHC class II genes in *R. omeimontis*.

Our results suggest that balancing selection plays an important role in the maintenance of polymorphisms of MHC class IIB genes in *R. omeimontis*. First, we found that the divergence in the amino acid sequence of MHC class IIB genes in *R. omeimontis* was greater than that in the nucleotide sequence, suggestive of positive selection. Second, positive selection signals were detected using several programs, although the deduced sites under selection were different between these programs. Three, six, and ten sites were inferred to be positively selected using the FEL, MEME (Pond and Frost, 2005) and codeml (Yang, 1997; Yang, 2007) methods, respectively. Only four of the twelve sites were detected more than once (Table 3). This difference in the deduced sites could be caused by not only different levels of power of detecting selection for each program, but also the different selection types that were detected. In this study, MEME identified three more positively selected sites than FEL, and all three of these sites overlapped with a putative ABS. Given that the FEL method detected sites under constant selective pressure and that MEME identified individual

Table 1 Sequences of the newly designed primers used in this study.

Primer	Sequence (5'-3')
MHCIIB-4F	GGABACANTCTGCARACACAACTA
MHCIIB-4R	CTTRYTBYKDGCDGATTCAGA
II 5-NGSP	TGTCCACAAAGCAAGTCAGG
RB-0a	TATTACCGGAACGGGACGGACGATA
RB-1a	ACGATGGACTCCAGTCCGGCCCGGACG ATATCAGGTTTCTGCAACG
RB-1ac	CGGACGATATCAGGTTTCTGCAACG
RB-2a	TACTGGAACAGCCAGCCCGATGTATTAG
SP1a	GGAAGGGTCAGGAAGGGTAAGAGA
SP1b	CCGGTACTCACATTTGCGGTCTA
SP2	GCCCGTACTCTGGCTAATACATCG
SP3	TCGCTGTCGAAGTACACATACTCCTC
MHCIIB Intron1-F	CGTGTGATGGTGCAGTGACCT
MHCIIB Intron1-F01	TGTGTGTTGTGTTCTCTCCCTG
MHCIIB Intron2-R01	GGTCAGGAAGGGTAAGAGAGG

Table 2 The synonymous (dS) and non-synonymous (dN) nucleotide substitution rates in different regions of exon 2 of the newly identified MHC class II alleles.

Region	dN (SE)	dS (SE)	dN/dS	Stat.	*P*
Entire exon 2	0.042 (0.008)	0.031 (0.011)	1.355	0.889	0.188
ABS	0.149 (0.038)	0.086 (0.075)	1.733	0.816	0.208
Non-ABS	0.024 (0.006)	0.021 (0.007)	1.143	0.421	0.337

The dS and dN values were computed according to the Nei-Gojobori method (Nei and Gojobori, 1986), and the standard errors (SE) values, which are shown in parentheses, were obtained using 1000 bootstrap replicates. Stat. and *P* represent the test statistic and the *P*-value, respectively, according to the codon-based test of positive selection.

Table 3 Summary of the codon sites undergoing positive selection identified using different methods.

Method	Amino acid sites																	
	4	6	8	21	23	25	32	41	42	52	55	62	65	66	69	73	79	81
ABS	+	+	+		+	+	+		+	+	+	+	+	+	+	+		+
PAML		+		+		+	+	+	+				+		+		+	+
FEL		+		+														+
MEME		+		+						+			+			+		+

The numbers for the amino acid sites correspond to the alignment shown in Figure 2.

sites under both episodic and pervasive positive selection (Murrell *et al*., 2012), these three additional sites were probably under episodic selection. Because the pathogens that organisms encounter commonly change over time, episodic selection is likely to become the primary selection type associated with MHC genes. Moreover, 5 of the 6 sites identified by MEME and 7 of the 10 sites identified by codeml were on an ABS, further suggesting that pathogen-mediated balancing selection contributed to the maintenance of the polymorphisms in MHC class II genes. Among the few deduced selected sites that were not located on an ABS, site 21 was notable because it was detected by all the three programs. Such inconsistent results reveal the slight inaccuracy of deducing ABSs in MHC IIB alleles of *R. omeimontis* from human HLA sequences, probably because of differences in the crystal structure of MHC class II molecules between treefrogs and humans. The misjudgment of ABSs might have partially contributed to the non-significant results of the global selection test examining the ABSs of MHC class II genes in *R. omeimontis*. Previous research has demonstrated that the detection of selection would be hampered when alleles from multiple loci are included in the analysis (Babik *et al*., 2008, 2009; May and Beebee, 2009). Therefore, efforts should be made in future studies to design locus-specific primers and isolate a single locus of MHC genes (Spurgin and Richardson, 2010; Bataille *et al.*, 2015; Lillie *et al.*, 2015).

5. Conclusion

This study presents the first characterization of MHC IIB alleles in *R. omeimontis*. New primers were successfully designed to isolate the full-length MHC IIB exon 2 sequences in *R. omeimontis*. We identified a total of 102 alleles in 146 *R. omeimontis* individuals. At least 4 MHC class IIB loci were found in *R. omeimontis*, but the allelic polymorphism was not high. No significant gene recombination was detected in the newly isolated alleles. In contrast, gene duplication and pathogen-mediated balancing selection likely played an important role in the evolution of MHC IIB genes in *R. omeimontis*. This study provides information to the future comprehensive researches including the evolution of amphibian MHC genes and the role of MHC polymorphisms in disease resistance for amphibians in China.

Acknowledgements This study was supported by the National Natural Science Foundation of China (No. 31201713, No. 31270425 and No. 31470442).

References

Aguilar A., Roemer G., Debenham S., Binns M., Garcelon D., Wayne R. K. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. Proc Natl Acad Sci U S A, 101: 3490–3494

Andersson L., Mikko S. 1995. Generation of MHC class II diversity by intra- and intergenic recombination. Immunol Rev, 143: 5–12

Babik W., Pabijan M., Arntzen J. W., Cogalniceanu D., Durka W., Radwan J. 2009. Long-term survival of a urodele amphibian despite depleted major histocompatibility complex variation. Mol Ecol, 18: 769–781

Babik W., Pabijan M., Radwan J. 2008. Contrasting patterns of variation in MHC loci in the Alpine newt. Mol Ecol, 17: 2339–2355

Bai C., Garner T. W., Li Y. 2010. First evidence of *Batrachochytrium dendrobatidis* in China: discovery of chytridiomycosis in introduced American bullfrogs and native amphibians in the Yunnan Province, China. Ecohealth, 7: 127–134

Baker C. S., Vant M. D., Dalebout M. L., Lento G. M., O'Brien S. J., Yuhki N. 2006. Diversity and duplication of DQB and DRB-like genes of the MHC in baleen whales (suborder: Mysticeti). Immunogenetics, 58: 283–296

Balakrishnan C., Ekblom R., Volker M., et al. 2010. Gene duplication and fragmentation in the zebra finch major histocompatibility complex. BMC Biol, 8: 29

Baláž V., Vörös J., Civiš P., et al. 2014. Assessing risk and guidance on monitoring of *Batrachochytrium dendrobatidis* in Europe through identification of taxonomic selectivity of infection. Conserv Biol, 28: 213–223

Bataille A., Cashins S. D., Grogan L., et al. 2015. Susceptibility of amphibians to chytridiomycosis is associated with MHC class II conformation. Proc Roy Soc B: Biol Sci, 282

Begovich A. B., McClure G. R., Suraj V. C., Helmuth R. C., Fildes N., Bugawan T.L., Erlich H.A., Klitz W. 1992. Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. J Immunol, 148: 1249–1258

Berger L., Speare R., Daszak P., et al. 1998. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. Proc Natl Acad Sci U S A, 95: 9031–9036

Bernatchez L., Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? J Evol Biol, 16: 363–377

Bollmer J. L., Dunn P. O., Whittingham L. A., Wimpee C. 2010. Extensive MHC class II B gene duplication in a passerine, the common Yellowthroat (*Geothlypis trichas*). J Hered, 101: 448–460

Bos D. H., DeWoody J. A. 2005. Molecular characterization of major histocompatibility complex class II alleles in wild tiger salamanders (*Ambystoma tigrinum*). Immunogenetics, 57: 775–781

Brown J. H., Jardetzky T. S., Gorga J. C., et al. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature, 364: 33–39

Bryja J., Galan M., Charbonnel N., Cosson J. F. 2006. Duplication, balancing selection and trans-species evolution explain the high levels of polymorphism of the DQA MHC class II gene in voles (Arvicolinae). Immunogenetics, 58: 191–202

Edwards S. V., Hedrick P. W. 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. Trends Ecol Evol, 13: 305–311

Fei L., Ye C. Y. 2001. The colour handbook of amphibians of Sichuan. China Forestry Publishing House, Beijing (China)

Garner T. W., Perkins M. W., Govindarajulu P., et al. 2006. The emerging amphibian pathogen *Batrachochytrium dendrobatidis* globally infects introduced populations of the North American bullfrog, *Rana catesbeiana*. Biol Lett, 2: 455–459

Garrigan D., Hedrick P. W. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution, 57: 1707–1722

Goldman N., Yang Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol, 11: 725–736

Hauswaldt J. S., Stuckas H., Pfautsch S., Tiedemann R. 2007. Molecular characterization of MHC class II in a nonmodel anuran species, the fire-bellied toad *Bombina bombina*. Immunogenetics, 59: 479–491

Hughes A. L., Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci U S A, 86: 958

Hughes A. L., Nei M. 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol Biol Evol, 7: 491–514

Hughes A. L., Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet, 32: 415–435

Jakobsen I. B., Wilson S. R., Easteal S. 1998. Patterns of reticulate evolution for the classical class I and II HLA loci. Immunogenetics, 48: 312–323

Jeffery K. J. M., Bangham C. R. M. 2000. Do infectious diseases drive MHC diversity? Microbes Infect, 2: 1335–1341

Kamath P. L., Getz W. M. 2011. Adaptive molecular evolution of the major histocompatibility complex genes, DRA and DQA, in the genus Equus. BMC Evol Biol, 11: 128

Kiemnec-Tyburczy K. M., Richmond J. Q., Savage A. E., Lips K. R., Zamudio K. R. 2012. Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. Heredity, 109: 146–155

Kiemnec-Tyburczy K. M., Richmond J. Q., Savage A. E., Zamudio K. R. 2010. Selection, trans-species polymorphism, and locus identification of major histocompatibility complex class IIbeta alleles of New World ranid frogs. Immunogenetics, 62: 741–751

Laurens V., Chapusot C., del Rosario Ordonez M., Bentrari F., Padros M. R., Tournefier A. 2001. Axolotl MHC class II beta chain: predominance of one allele and alternative splicing of the β1 domain. Eur J Immunol, 31: 506–515

Librado P., Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25: 1451–1452

Li F., Shu Y. L., Wu H. L. 2012. Polymorphism of exon 2 of MHC Class II B gene in the Chinese concave-eared torrent frog (*Odorrana tormota*). Biodivers Sci, 20: 184–192

Lillie M., Grueber C. E., Sutton J. T., et al. 2015. Selection on MHC class II supertypes in the New Zealand endemic Hochstetter's frog. BMC Evol Biol, 15: 63

Lips K. R., Brem F., Brenes R., et al. 2006. Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. Proc Natl Acad Sci U S A, 103: 3165–3170

Liu Y. G., Chen Y. 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. Biotechniques, 43: 649–656

Martin D. P., Murrell B., Golden M., Khoosal A., Muhire B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol, 1: vev003

May S., Beebee T. J. C. 2009. Characterisation of major histocompatibility complex class II alleles in the natterjack toad, *Bufo calamita*. Conserv Genet Resour, 1: 415–417

May S., Zeisset I., Beebee T. J. C. 2011. Larval fitness and immunogenetic diversity in chytrid-infected and uninfected natterjack toad (*Bufo calamita*) populations. Conserv Genet, 12: 805–811

Morgan J. A., Vredenburg V. T., Rachowicz L. J., et al. 2007. Population genetics of the frog-killing fungus *Batrachochytrium dendrobatidis*. Proc Natl Acad Sci U S A, 104: 13845–13850

Murrell B., Wertheim J. O., Moola S., Weighill T., Scheffler K., Kosakovsky Pond S. L. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet, 8: e1002764

Nei M., Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol, 3: 418–426

Nei M., Gu X., Sitnikova T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci U S A, 94: 7799–7806

Nei M., Rooney A. P. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet, 39: 121–152

O'Brien S. J., Evermann J. F. 1988. Interactive influence of infectious disease and genetic diversity in natural populations. Trends Ecol Evol, 3: 254–259

Ohta Y., Goetz W., Hossain M. Z., Nonaka M., Flajnik M. F. 2006. Ancestral organization of the MHC revealed in the amphibian Xenopus. J Immunol, 176: 3674–3685

Piertney S. B., Oliver M. K. 2006. The evolutionary ecology of the major histocompatibility complex. Heredity, 96: 7–21

Pond S. L. K., Frost S. D. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol, 22: 1208–1222

Pounds J. A., Bustamante M. R., Coloma L. A., et al. 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. Nature, 439: 161–167

Reusch T. H., Langefors Å. 2005. Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback *Gasterosteus aculeatus*. J Mol Evol, 61: 531–541

Rozen S., Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol, 132: 365–386

Sato A., Tichy H., Grant P. R., Grant B. R., Sato T., O'HUigin C. 2011. Spectrum of MHC class II variability in Darwin's finches and their close relatives. Mol Biol Evol, 28: 1943–1956

Sato K., Flajnik M. F., Du Pasquier L., Katagiri M., Kasahara M. 1993. Evolution of the MHC: isolation of class II beta-chain cDNA clones from the amphibian *Xenopus laevis*. J Immunol, 150: 2831–2843

Savage A. E., Zamudio K. R. 2011. MHC genotypes associate with resistance to a frog-killing fungus. Proc Natl Acad Sci U S A, 108: 16705–16710

Schaschl H., Wandeler P., Suchentrunk F., Obexer-Ruff G., Goodman S. J. 2006. Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. Heredity, 97: 427–437

She J. X., Boehme S. A., Wang T. W., Bonhomme F., Wakeland E. K. 1991. Amplification of major histocompatibility complex class II gene diversity by intraexonic recombination. Proc Natl Acad Sci U S A, 88: 453–457

Shu Y. L., Hong P., Yang Y. W., Wu H. L. 2013. An endemic frog harbors multiple expression loci with different patterns of variation in the MHC class II B gene. J Exp Zool Part B: Mol Develop Evol, 9999: 1–10

Siddle H. V., Kreiss A., Eldridge M. D. B., et al. 2007. Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci U S A, 104: 16221–16226

Smith S., de Bellocq J. G., Suchentrunk F., Schaschl H. 2011. Evolutionary genetics of MHC class II beta genes in the brown hare, *Lepus europaeus*. Immunogenetics, 63: 743–751

Sommer S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front Zool, 2: 16

Spurgin L. G., Richardson D. S. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc Roy Soc B: Biol Sci, 277: 979–988

Stuart S. N., Chanson J. S., Cox N. A., et al. 2004. Status and trends of amphibian declines and extinctions worldwide. Science, 306: 1783–1786

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol, 30: 2725–2729

Tong J. C., Bramson J., Kanduc D., Chow S., Sinha A. A., Ranganathan S. 2006. Modeling the bound conformation of *Pemphigus vulgaris*-associated peptides to MHC Class II DR and DQ alleles. Immunol Res, 2: 1

Wei Y., Xu K., Zhu D.Z., Chen X.F., Wang X.L. 2010. Early-spring survey for *Batrachochytrium dendrobatidis* in wild *Rana dybowskii* in Heilongjiang Province, China. Dis Aquat Organ, 92: 241–244

Xu S., Ren W., Zhou X., Zhou K., Yang G. 2010. Sequence polymorphism and geographical variation at a positively selected MHC-DRB gene in the finless porpoise (*Neophocaena phocaenoides*): implication for recent differentiation of the Yangtze finless porpoise? J Mol Evol, 71: 6–22

Xu T. J., Sun Y. N., Wang R. X. 2011. Allelic polymorphism, gene duplication and balancing selection of the MHC class II DAB gene of *Cynoglossus semilaevis* (Cynoglossidae). Genet Mol Res, 10: 53–64

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci, 13: 555–556

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol, 24: 1586–1591

Yang Z., Nielsen R., Goldman N., Pedersen A. M. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics, 155: 431–449

Yang Z., Wong W. S., Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol, 22: 1107–1118

Yu X., Zheng R., Zhang J., Shen B., Dong B. 2014. Genetic polymorphism of major histocompatibility complex class IIB alleles and pathogen resistance in the giant spiny frog *Quasipaa spinosa*. Infect Genet Evol, 28: 175–182

Zeisset I., Beebee T. C. 2013. Bufo MHC class II loci with conserved introns flanking exon 2: cross-species amplification with common primers. Conserv Genet Resour, 5: 211–213

Zeisset I., Beebee T. J. 2009. Molecular characterization of major histocompatibility complex class II alleles in the common frog, *Rana temporaria*. Mol Ecol Resour, 9: 738–745

Zhao E. M., Adler K. 1993. Herpetology of China. Society for the Study of Amphibians and Reptiles, Oxford (Ohio)

Zhao M., Wang Y., Shen H., et al. 2013. Evolution by selection, recombination, and gene duplication in MHC class I genes of two Rhacophoridae species. BMC Evol Biol, 13: 113

Zhu R., Chen Z. Y., Wang J., et al. 2013. Extensive diversification of MHC in Chinese giant salamanders *Andrias davidianus* (Anda-MHC) reveals novel splice variants. Dev Comp Immunol, 42: 311–322

Table S1 Selected codons estimated using the FEL method.

Codon	dS	dN	dN/dS	Normalized dN-dS	*P*-value
6	7.891E–17	8.141	1.032E+17	10.835	0.007
21	5.449E–15	5.660	1.039E+15	7.532	0.010
81	1.385E–16	10.370	7.489E+16	13.801	0.042
60	5.620	0	0	–7.479	0.024

Only those sites displaying a *P*-value <0.05 are listed in this table. The normalized dN-dS was computed by dividing dN-dS by the total length of the appropriate tree.

Table S2 Positively selected codons estimated using the MEME method.

Codon	α	β–	Pr[β= β–]	β+	Pr[β= β+]	*P*-value	*q*-value
6	0	0	0.185	15.889	0.815	0.002	0.067
21	0	0	0.751	29.925	0.249	0.002	0.051
52	0	0	0.949	140.390	0.051	7.138E–05	0.006
65	0	0	0.964	10000	0.036	0.000	0.013
73	0	0	0.910	34.267	0.090	0.030	0.451
81	0	0	0.140	17.594	0.860	0.023	0.401

This table shows the distribution of non-synonymous (β) and synonymous (α) substitution rates for positively selected sites inferred using the MEME model. Only those sites displaying a *P*-value < 0.05 are listed.

Table S3 Evaluation of the goodness of fit for different models of codon evolution and the likelihood ratio tests comparing the nested models.

Model code	*f*	LnL	*P*arameter estimates	2ΔL (*P*)	*P*ositively selected sites
M1a (neutral)	2	–2375.941	*K*=0.874 *p*0=0.837 *p*1=0.163 *ω*0=0.089 *ω*1=1.000		Not allowed
M2a (selection)	4	–2153.966	*K*=2.080 *p*0=0.341 *p*1=0.558 *p*2=0.101 *ω*0=0.275 *ω*1=1.000 *ω*2=25.046	443.948 (*P*<0.001)	**6**, **21**, **25**, *32*, **41**, **42**, **65**, **69**, **79**, **81**
M7 (β)	2	–2403.458	*K*=0.898 *p*=0.267 *q*=0.543		Not allowed
M8 (β and *ω*)	4	–2154.651	*K*=2.095 *p*=0.488 *q*=0.138 *p*0=0.899 *p*1=0.101 *ω*=26.695	497.613 (*P*<0.001)	**6**, 8 **21**, 23, **25**, **32**, **41**, **42**, **65**, 66, **69**, **79**, **81**, 86

ω refers to the ratio of nonsynonymous to synonymous substitution rates (dN/dS); *f* refers to the number of free parameters in the *ω* distribution; LnL refers to the log-likelihood; *p*n is the proportion of sites categorized in the *ω*n site class; and *p* and *q* are the shape parameters of the β function (for models M7 and M8). 2ΔL, the test statistic of the likelihood ratio test, indicates the double difference in log likelihood between the nested models (M2a vs. M1a and M8 vs. M7). Positively selected sites were identified in models M2a and M8 using the Bayes empirical Bayes procedure (Yang *et al.* 2005). Sites that were deduced to be under positive selection at the 95% or 99% confidence interval level are presented in italics and in bold, respectively.